

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets

(11)

EP 1 013 762 A2



(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
28.06.2000 Bulletin 2000/26

(51) Int. Cl.⁷: C12N 15/12, C07K 14/705,
C12Q 1/68

(21) Application number: 99126049.8

(22) Date of filing: 10.12.1993

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL
PT SE

(30) Priority: 11.12.1992 US 989537

(62) Document number(s) of the earlier application(s) in
accordance with Art. 76 EPC:
94903572.9 / 0 673 390

(71) Applicant:
THE GOVERNMENT OF THE UNITED STATES OF
AMERICA, as represented by THE SECRETARY
OF THE DEPARTMENT OF HEALTH AND HUMAN
SERVICES
Rockville, MD 20852-3804 (US)

(72) Inventor: Laborda, Jorge
Silver Spring, Maryland 20906 (US)

(74) Representative:
Mercer, Christopher Paul
Carpmaels & Ransford
43, Bloomsbury Square
London WC1A 2RA (GB)

Remarks:

This application was filed on 28 - 12 - 1999 as a
divisional application to the application mentioned
under INID code 62.

(54) **Delta-like gene expressed in neuroendocrine tumors**

(57) A polynucleotide molecule *dlk* is expressed in
neuroendocrine tumors, including small cell lung carci-
noma. A Dlk polypeptide encoded by *dlk* polynucleotide
molecule can be used in detecting the existence of a pri-
mary or secondary neuroendocrine tumor. Monoclonal
antibodies are produced against Dlk which are useful
for detection and therapy of a neuroendocrine tumor.

Description**Background of the Invention**

5 [0001] The expression of genes during the development of a pluripotent or progenitor cell into a differentiated, mature cell can provide a context for the study of tumorigenic cells whose origin is derived from such progenitor cells. In certain hematopoietic or epithelial tumors, malignant gene expression correlates substantially with the expression observed during normal development of the tissue from which the tumor originates, Gordon et al., *J. Cell Biol.* **108**: 1187 (1989); Godal et al., *Adv. Cancer Res.* **36**: 211 (1982). In fact, many biological activities of progenitor cells, including 10 cellular migration and tissue remodeling, resemble pathological activities of cancer cells, such as metastases and tumor invasion.

[0002] Neuroblastoma, a tumor of the adrenal gland which afflicts persons during early childhood, is another system in which tumor biology correlates with that of normal differentiation and morphogenesis of its progenitor cells (neuroblast). Neuroblastoma is an embryonal tumor that exhibits both undifferentiated and differentiated histopathology. The 15 development of neuroblastoma tumors mimics stages identifiable during histogenesis of its tissue of origin, the adrenal medulla. Cooper et al., *Cell Growth and Diff.* **1**: 149 (1989).

[0003] During the development of human adrenal medulla neuroblasts into mature chromaffin cells, four individual genes are expressed in a sequential pattern. Once a neuroblast is induced to differentiate along a neuroendocrine pathway, the progressive stages of chromaffin maturation are marked by a temporal expression of genes denoted TH, CGA, 20 pG2 and B2M (Cooper, *supra*. at page 153). Cooper found that the pattern of gene expression of these four markers in neuroblastoma cells mimics that of normal adrenal neuroblast arrested during three different stages of development.

[0004] One of these marker genes, pG2, was identified first in pheochromocytoma, a tumor of the adult adrenal medulla (Helman et al., *PNAS USA* **84**: 2336 (1987)). Helman reported that pG2 also is highly expressed normal 25 human adrenal cells.

[0005] Helman isolated a full-length cDNA from a human adrenal cDNA library, and identified a corresponding pG2 protein containing 286 amino acids, having a predicted molecular weight of 30.6 kilodaltons (kDa) (Helman et al., *Nucleic Acids Res.* **18**(3): 685 (1990)).

[0006] A gene having developmentally-regulated expression, paralleling that of pG2, would be useful for detecting 30 pheochromocytoma or neuroblastoma by genetic methods, especially since pG2 expression is restricted to the adrenal gland in non-malignant tissue.

Summary of the Invention

[0007] It is therefore an object of the present invention to provide a newly-isolated polynucleotide molecule, *dk*, which can be employed in genetic assays to provide a method for detection of a primary or secondary pheochromocytoma or neuroblastoma, or identification of a stage of these tumors.

[0008] It is also an object of the present invention to provide a method for detecting primary or secondary small cell lung carcinoma (hereafter, SCLC) or for staging tumor progression of SCLC, which employs *dk* polynucleotide molecules in genetic assays.

[0009] It is a further object to provide a polynucleotide molecule, designated *dk*, which encodes a corresponding Dlk polypeptide. Dlk polypeptides are useful for generating monoclonal or polyclonal antibodies having specificity for an epitope of the Dlk polypeptide.

[0010] Dlk-specific antibodies, and in particular, labeled monoclonal Dlk-specific antibodies, are useful for detection of primary or secondary neuroendocrine tumors. According to the present invention, Dlk-specific monoclonal antibodies 45 conjugated to a toxin are useful for treatment of primary or secondary neuroendocrine tumors, as well.

[0011] In accomplishing these and other objects of the invention, there has been provided, in accordance with one aspect of the present invention, an isolated polynucleotide molecule comprising a DNA sequence encoding a Dlk polypeptide.

[0012] An object of the present invention is to provide an isolated Dlk polypeptide consisting essentially of the 50 amino acid sequence shown in Figure 1B, or in Figure 1A.

[0013] Another object of the present invention is to provide an isolated polynucleotide molecule which encodes a human or mouse Dlk polypeptide consisting essentially of the amino acid sequence shown in Figures 1B, or 1A, respectively.

[0014] A further object of the invention is to provide a method for detecting a tumor which expresses *dk*, including 55 the steps of contacting RNA from a sample suspected of being tumorigenic with a *dk* polynucleotide molecule, under conditions permissive of hybridization between *dk* polynucleotide molecule and the sample, and detecting hybridization between the polynucleotide molecule and sample.

[0015] Yet another object of the invention is to provide a method for detecting a small cell lung carcinoma, including

the steps of contacting RNA from a sample of bronchial epithelial cells suspected of being tumorigenic with *dlk* polynucleotide molecules, under conditions permissive of hybridization between *dlk* polynucleotide molecules and the sample, and detecting hybridization between the polynucleotide molecules and sample.

5 Brief Description of the Drawings

[0016]

10 Figure 1 shows an alignment of mouse (Figure 1A) and human (Figure 1B) Dlk amino acid sequences. Identical amino acids are shown by the character (|). Similar amino acids are indicated by (^) and classified into the following groups: A, S & T; D & E; N & Q; R & K; I, L, M & V; and F, Y & W. Potential biologically significant sites, found in the database PROSITE (accessible commercially through Intelligenetics Inc. (Mountain View, CA)), are indicated by numbers: 1. N-glycosylation site; 2. Protein Kinase C phosphorylation site; 3. N-myristylation site; 4. Aspartic acid and asparagine hydroxylation site. Potential sites of cleavage in the signal peptide are indicated by (*).

15 Figure 2 shows the human *dlk* DNA sequence.

Figure 3 shows the mouse *dlk* DNA sequence.

20 Figure 4 shows an alignment of a consensus sequence of *dlk* EGF-like repeats with EGF-repeats found in several invertebrate homeotic genes. As described in Example 4, a *dlk* EGF-like repeat consensus sequence was obtained by alignment of 12 EGF-like repeats of *dlk* from both human and mouse. This consensus sequence then was aligned with the consensus sequences of several invertebrate homeotic genes (similarly obtained) and mouse EGF.

Detailed Description of the Preferred Embodiments

25 [0017] A human polynucleotide molecule, *dlk*, and a corresponding human polypeptide, Dlk, encoded by *dlk*, were discovered, isolated and characterized. The human *dlk* polynucleotide molecule was found to be expressed in pheochromocytoma, neuroblastoma, and SCLC tumors.

30 [0018] Dlk protein is about 383 amino acids in length and has a molecular weight of about 42 kDa. In addition to human *dlk*, other polynucleotide molecules belonging to the *dlk* family are provided according to the invention, including murine *dlk* (Figure 3) and a human variant-*dlk*, isolated from placenta as described herein.

35 [0019] According to the present invention, isolated polynucleotide molecules or fragments thereof belonging to the *dlk* family are useful in the detection of SCLC and neuroendocrine cancers. The expression patterns of *dlk* can be exploited both (1) to detect primary or secondary tumor cells by the presence of *dlk* and (2) to diagnose the stage of a tumor that expresses *dlk*, by measuring the level of *dlk* expression.

40 [0020] Dlk is a transmembrane protein having an expression pattern, in normal non-fetal tissues, which is restricted to the adrenal gland. As a consequence, Dlk is a readily accessible target for antibody imaging or therapy of SCLC, pheochromocytoma and neuroblastoma tumors. According to the present invention, antibodies having specificity for the Dlk protein are made and employed to detect or treat cells which produce the Dlk protein.

45 [0021] Human *dlk* cDNA comprises a polynucleotide molecule having the sequence shown in Figure 2, as determined by nucleotide sequence analysis. The open reading frame, nucleotides 174(ATG) - 1322(TAA), is 1149 nucleotides long. The mouse *dlk* polynucleotide molecule comprises a DNA sequence having an open reading frame, nucleotides 134(ATG) - 1288 (TAA), of 1155 nucleotides, as shown in Figure 3. The murine Dlk protein is about 385 amino acids and has a molecular weight of about 42 kDa.

50 [0022] According to the present invention, a variant of human Dlk is identified in which an amino acid is deleted. A cDNA encoding "variant-Dlk," in which amino acid number 347 of the amino acid sequence shown, in Figure 1(B) is deleted, was isolated from a total human placental cDNA library. Amino acid position 347 occurs in an intracellular domain of the protein. The placental library containing variant-*dlk* also contained substantial amounts of the non-variant form, that is, *dlk* polynucleotide molecule shown in Figure 1B.

55 [0023] The *dlk* polynucleotide molecule was identified by examination of cDNA expression products of human SCLC (hSCLC) lines which were responsive to stimulation with the ligand, gastrin-releasing peptide (GRP), a neuropeptide implicated in the release of gastrin through its interaction with a G-protein-coupled receptor, GRP receptor. GRP (peptide) is a mitogen for both normal lung epithelial cell, and SCLC cells, and for murine Swiss 3T3 fibroblasts.

[0024] GRP-responsive hSCLC lines were compared with murine fibroblast cell lines that were differentially responsive to GRP. This approach, as detailed in Example 1, yielded a partial length cDNA molecule that hybridized a 1.6 Kb mRNA expressed both in responsive fibroblasts and responsive SCLC lines. A commercially available Swiss 3T3 fibroblast cDNA library was screened with the partial length cDNA, which yielded several clones having 1.6 Kb inserts, which then were sequenced.

[0025] A computer search of the databases "Swissprot" and "NBRF Protein," described by Devereux et al., *Nuc.*

Acids Res. 12(1): 387 (1984), indicated a high degree of homology between Dlk and proteins encoded by several homeotic genes, identified in Example 3. Homeotic genes are development-controlling regulatory genes that assign spatial identity to groups of cells with respect to their morphogenic fates. In segmented organisms, for example, homeotic genes are required for the proper morphogenesis of a distinct region (such as a leg, or antennae) and act by controlling the activities of other genes during development. The Dlk protein of the present invention exhibited highest homology with the protein Delta, a neurogenic locus involved in normal neural differentiation in *Drosophila*. Thus, the present protein was designated "Dlk" for being "delta-like."

[0026] Mouse and human Dlk protein sequences share 86.2% identity as well as many potential sites of biological importance, including 6 epidermal growth factor (EGF)-like repeats, a transmembrane domain, and a signal peptide domain at the amino terminus. Based upon these structural features, *dlk* appears to be a new member of the family of EGF-like neurogenic genes of *Drosophila*, which are involved in developmental decisions of the embryonal ectoderm to differentiate into epidermal or neuronal cells.

[0027] The expression pattern of *dlk* and its sequence homology with homeotic proteins support the notion that *dlk* functions in the differentiation pathway for cells of the chromaffin lineage. As detailed in Example 2, *dlk* is expressed in primary and secondary pheochromocytoma and neuroblastoma cells, and in normal (nonhistopathological) human adrenal medulla and placental cells. According to the present invention, SCLC and neuroblastoma are the only tumors known to express *dlk* as a function of differentiation.

[0028] An isolated *dlek*, *dlek*-variant, and murine *dlek* polynucleotide and protein products are employed in diagnostic methods (described further below) and are made according to the following description. Hereafter, the techniques and applications described for *dlek* polynucleotide molecule (DNA, RNA) and Dlk protein are intended to be useful for DNA, RNA and protein of murine *dlek*, and of variant-*dlek*, as well.

[0029] A Dlk polypeptide, according to the present invention, is produced by recombinant DNA techniques, such as those described by Maniatis et al., MOLECULAR CLONING - A LABORATORY MANUAL, Cold Spring Harbor Laboratory (1982). Methods specifically applicable to cloning the *dlek* polynucleotide molecule are described in Example 1.

[0030] The *dlek* polynucleotide molecule of Figure 1B can be cloned into suitable expression vectors and expressed in prokaryotic, insect or eukaryotic expression systems, including *Baculovirus* or *E. coli* (Boehringer Manhein). Using conventional techniques, therefore, a polynucleotide sequence encoding a Dlk protein, can be obtained as a cDNA from mRNA from a commercial adrenal medulla or Swiss 3T3 fibroblast library, or from SCLC, neuroblastoma or pheochromocytoma cell lines. The mRNA can be converted to double-stranded DNA using cDNA cloning techniques well-known to the art, including PCR-based techniques. Linkers or tails may be added to the ends of the double-stranded DNA to provide convenient restriction sites. After restriction enzyme digestion, the DNA may be introduced into a cloning vector, such as a plasmid, which has been digested with a restriction enzyme that generates appropriately compatible ends. A suitable plasmid vector in this context is pGEX-λ (Pharmacia). Following ligation, by means of standard techniques, the DNA is introduced into a cell, where its expression produces the desired protein.

[0031] Alternatively, a Dlk polypeptide is produced using a commercially available *in vitro* translation kit from NEN (Boston, MA), as detailed in Example 1. This kit employs a translation system (including ribosomes, polymerases, amino acids, etc.) derived from rabbit reticulocyte lysates.

[0032] The term "isolated," with reference to the *dlek* polynucleotide molecule indicates that such a molecule is free of the proteins with which it is normally associated, such as histones. An isolated form of the *dlek* is substantially free of other DNA that does not function to regulate, promote, enhance or otherwise modulate its expression.

[0033] The term "isolated," with reference to Dlk protein, connotes a polypeptide that is free of other proteins with which it is normally associated.

[0034] An isolated *dlek* polynucleotide molecule is useful for detecting primary SCLC and the metastatic spread of SCLC and other neuroendocrine cancers. More specifically, the present invention provides a method of tumor detection which includes the steps of contacting a sample suspected of containing a tumor with a *dlek* polynucleotide molecule, and detecting expression of *dlek* polynucleotide products (DNA, RNA, mRNA) in non-adrenal cells. Detection of a *dlek* polynucleotide product identifies the cells as metastatic cells (secondary tumor) of neuroblastoma, pheochromocytoma or SCLC, or as a primary tumor of SCLC.

[0035] The ability to detect *dlek*-expressing cells is useful for both tumor detection and tumor identification. After detecting *dlek*-expression, a tumor type is determined by detecting either a tumor-specific marker, tumor-specific morphology, or by presentation by the patient of a clinical pathology that is distinctly associated with any of the tumors selected from the group including neuroblastoma, pheochromocytoma or SCLC. For example, information such as the identification of a cellular marker, histological feature or disease symptom which is specific to one of the tumors of neuroblastoma, pheochromocytoma or SCLC, is recognized.

[0036] If *dlek* expression is detected in cells of a sample taken from bronchial epithelial tissue or tissue removed from the lung, the observation identifies a primary SCLC. It is preferred that a second step of confirming the origin of the detected *dlek*-expressing tumor cells as SCLC be performed by detection of a marker, histological feature, or presentation of a distinctive symptom associated with this tumor. For example the histology of an "oat cell" commonly associated

with SCLC is detected to confirm the presence of SCLC.

[0037] The expression of *dlk* is detected by hybridization with a *dlk* polynucleotide probe. This method includes the steps of contacting a suspected tumor sample with a *dlk* polynucleotide molecule, and detecting the hybridization between the polynucleotide molecule and sample. A positive hybridization signal indicates that the sample is of tumor origin.

[0038] The polynucleotide molecule or "dlk probe" used to detect *dlk* expression is a labeled fragment of *dlk*, or preferably a full-length *dlk* DNA molecule which will hybridize to mRNA or DNA from normal adrenal and neuroendocrine tumor cells. Probes complementary to *dlk* are prepared by conventional methods, and are preferably allowed to hybridize to mRNA or DNA, using conventional *in situ* hybridization techniques. Unhybridized probe is removed by nuclease digestion.

[0039] *In situ* hybridization techniques which are known in the art may employ the use of fluorescent labels and radiolabels which can be easily quantitated by fluorescence microscopy or autoradiography, respectively. Generally, fluorescent labels will be preferred. Another labelling technique may employ enzymatic tags which generate readily quantifiable colorimetric or chemiluminescent signals. The hybridization intensity detected in their procedures reflects the amount of *dlk* within the biological sample.

[0040] The *dlk* polynucleotide of the invention can be employed as a probe in RNA ("Northern") blotting procedures. According to this method, RNA is first isolated from tissue by any of a number of standard procedures (Lehrach, H., *Biochemistry*, 16: 4743 (1975)). The RNA sample is then subjected to denaturing gel electrophoresis and is transferred to a nitrocellulose membrane or other solid support matrix. The *dlk* mRNA can be detected by hybridization of radioactively or non-radioactively labelled *dlk*, or *dlk* fragments, preferably under high stringency conditions as would be familiar to one of ordinary skill in the art. The amount of hybridization can be quantified by densitometry.

[0041] In yet another embodiment of the present invention, the polymerase chain reaction ("PCR") is used to detect *dlk* DNA or mRNA in a sample. To perform PCR, a pair of *dlk* sequence specific polynucleotide primers is employed, which hybridize to opposite strands of the *dlk* gene at offset positions on the double helix. Such primers, taken from the *dlk* polynucleotide sequences provided in accordance with the invention, represent fragments which preferably are unique to *dlk*, e.g., sequences having low homology with other proteins than Dlk. Two exemplary *dlk*-specific primer sequences useful in this context include the following sequences, which encode a portion of the intracellular region of Dlk:

30 5'-CAA GCC CGA GTT CAC AGG TC-3'
5'-TCG GGG AAG ATG TTG AC-3'.

Other such primer pairs can be selected and utilized, as well.

[0042] The primers provide initiation points for DNA synthesis. In the presence of DNA polymerase, the four deoxyribonucleotide triphosphates ("dNTPs") and other necessary cofactors, all of which are well known to the art, new DNA strands are synthesized complementary to the templates which hybridized with the primers. Several rounds of synthesis are carried out, with allowance for denaturation of the double stranded products between rounds. Preferably, a thermal stable DNA polymerase is used so that it is not necessary to add enzyme anew for each round of synthesis.

[0043] The PCR produces a double stranded DNA amplification product which has the same sequence as the original stretch of the *dlk* DNA defined by the ends of the primer pair sequences. The amount of PCR product indicates the amount of *dlk* DNA or *dlk* mRNA in the sample. The product can be detected by a variety of methods well-known in the art. The PCR products can be resolved by agarose or polyacrylamide electrophoresis and detected by fluorescence staining, such as ethidium bromide. Alternatively, one of the dNTPs may be labelled and the PCR products may be determined by measuring incorporation of the labelled dNTP. A variety of other methods for resolving, detecting and quantitating PCR products are well-known to those of ordinary skill in the art.

[0044] The PCR may be made specific for either *dlk* DNA or *dlk* mRNA. For instance, RNase or DNase may be used to remove one template or the other from the sample, and the use of primers that distinguish between the gene and the message, for example, a primer that hybridizes to a sequence in the untranscribed region of the promoter will be specific for the *dlk* gene, and not for the *dlk* mRNA.

[0045] Other techniques suitable to the claimed methods are readily apparent to the skilled artisan and can include Nuclease Protection Assays, ELISA and Western blotting. Several assay techniques which are based upon immunological reactions between antigens and antibodies are contemplated by the invention. In particular, assays which use antibodies having specificity for Dlk protein are useful to detect cells which produce Dlk protein.

[0046] Antibodies having specificity for Dlk-expressing cells are obtained by immunizing an animal with Dlk protein. In this context, the term "antibody" encompasses both monoclonal and polyclonal antibodies. Such an antibody can belong to any antibody class (IgG, IgM, IgA, etc.). According to the present invention, an entire Dlk polypeptide is injected into an animal for the purpose of obtaining polyclonal antibodies, or for obtaining lymphocytes or spleen cells for production of monoclonal antibodies.

[0047] The general techniques of monoclonal antibody (MAb) production, such as those described by Kohler and Milstein, *Nature* 256:495 (1975), are applied to produce a monoclonal antibody having specificity for Dlk protein. This procedure includes the steps of isolating lymphocytes of an animal which has been sensitized or injected with Dlk polypeptide, fusing them with a myeloma partner to produce hybridomas, then screening the hybridomas for production of "anti-Dlk antibodies" which exhibit binding specificity for a Dlk polypeptide.

[0048] The term "Antibody" also encompasses fragments, like FAb and F(AB¹)₂, of anti-Dlk antibodies, and conjugates of such fragments, and so-called "antigen binding proteins" (single-chain antibodies) which are based on anti-Dlk antibodies, in accordance, for example, with U.S. patent No. 4,704,692, the contents of which are hereby incorporated by reference. Alternatively, MAbs or fragments thereof can be produced via the expression of genes which encode variable regions of such an MAb in host cells like *E. coli*, see, e.g., Ward et al., *Nature* 341: 544-546 (1989), or transfected murine myeloma cells. See Verhoyen et al., *BioAssays* 8: 74 (1988); Gillies et al., *Biotechnol.* 7: 799-804 (1989); Nakatani et al., *Biotechnol.* 7: 805-10 (1989).

[0049] Assays in which the above antibodies are employed can include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays, immunoelectrophoresis, and the like. Also useful diagnostically are immunohistochemical techniques which employ monoclonal antibodies of known reactivity.

[0050] In accordance with this aspect of the present invention, a sample is obtained from a person to detect: (1) a small cell lung carcinoma, by removing a body fluid or tissue suspected of harboring a tumor, such as alveolar, bronchial, or respiratory epithelial cells obtained from a bronchial wash, nasopharyngeal aspirates, throat swabs or the like; (2) a metastasized neuroendocrine tumor, by biopsy, taken from tissue other than the adrenal gland (including cortex and medulla. Immuno-histochemical studies can be performed on such cells using a monoclonal antibody specific for Dlk.

[0051] Diagnostic applications of these antibodies are exemplified, according to the present invention, by the use of a kit containing an anti-Dlk antibody, which undergoes a reaction with a biological sample to detect Dlk protein. Such a reaction involves the binding of anti-Dlk antibody to Dlk antigen, under conditions permissive of binding. The observation of an antibody-antigen complex in a biological sample indicates a positive result. A kit of this sort could be used to detect the extent of expression of Dlk in a particular biological sample from an individual, animal, or cell line.

[0052] Such an immunodiagnostic kit can include anti-Dlk antibody and a receptacle for containing the antibody in a sterilized form. The kit can further include anti-isotype serum antibody which recognizes the anti-Dlk antibody (Fc portion) and which is conjugated to a label, such as an enzyme or fluorescent moiety.

[0053] In a preferred embodiment, a radiolabeled anti-Dlk antibody is provided. Such an antibody, preferably a monoclonal antibody, is administered to an animal or person for imaging purposes. After a suitable period of time for the administered antibodies to bind Dlk expressing cells, a gamma camera machine is applied to detect the presence of labeled antibodies within the organism. Such a procedure provides information as to where in the organism a primary or secondary Dlk-expressing neuroendocrine tumor is located.

[0054] A therapeutic application of anti-Dlk monoclonal antibodies includes administration of anti-Dlk immunotoxins. Conjugation of an anti-Dlk monoclonal antibody to a toxin, such as *Pseudomonas* exotoxin or other toxins commonly conjugated to an antibody by means of a conventional antibody-toxin linkage. Hertler et al., *J. Clin. Oncol.* 7(12): 1932 (1989), describe methodologies for creating an antibody-toxin linkage, and is incorporated by reference herein. Thus, the anti-Dlk monoclonal antibody-toxin conjugates described are administered to an individual to target and selectively kill Dlk-expressing cells present in neuroendocrine tumors.

[0055] Similarly, a kit is provided which contains anti-Dlk immunotoxins in a receptacle. A kit can include the anti-Dlk immunotoxins and a pharmaceutical excipient in a receptacle.

[0056] The present invention is further described with reference to the following, illustrative examples. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art of the invention. Although any method and materials similar or equivalent to those described herein can be used in the practice of the invention, the preferred methods and materials have been described. Unless mentioned otherwise, the techniques employed or contemplated herein are standard methodologies known to the art. The materials, methods and examples are illustrative only and not limiting.

50 **Example 1. IDENTIFICATION OF *dlk* POLYNUCLEOTIDE AND POLYPEPTIDE MOLECULES**

Identification of *dlk*

[0057] In investigating molecules associated with the gastrin-releasing peptide (GRP) responsive phenotype, genes were identified which both were: (1) expressed differentially between responsive murine Swiss and unresponsive murine Balb/c 3T3 fibroblasts, and (2) expressed in GRP-responsive human SCLC cell lines. The rationale for this approach was that gene products which correlated with a GRP-responsive phenotype would be missing from Balb/c and unresponsive SCLC cell lines, but present in Swiss 3T3 fibroblasts and responsive SCLC cell lines.

5 [0058] A differential cDNA library was constructed that was enriched for clones expressed in Swiss 3T3 but not in Balb/c 3T3 fibroblasts. The differential library of Swiss 3T3 compared with Balb/c 3T3 fibroblasts was constructed as explained in detail in Timblin et al., *Nucleic Acids Res.* **18**: 1587 (1990). The RNA isolation, electrophoresis, northern blots, and hybridization techniques were performed as described in Davis, et al., *BASIC METHODS IN MOLECULAR BIOLOGY*, Elsevier, New York, (1986). The nucleic acid probes were labelled with ^{32}P dCTP (Amersham, Arlington Heights, IL) by random priming as described in Ausubel et al., *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley and Sons, New York 3.5.9.-3.5.10 (1991).

10 [0059] A partial length clone (150 nucleotides long) isolated from this differential library hybridized with a 1.6 kilobase mRNA which exhibited an expression pattern consistent with the two screening requirements. This partial length clone was then used to screen a commercially available oligo dT-primed cDNA library of Swiss 3T3 fibroblasts in the λ ZAPII vector (Stratagene (La Jolla, CA)), to obtain a full length clone.

15 [0060] Screening procedures and plasmid rescue of positive λ ZAPII clones were performed following the manufacturer's (Stratagene) protocol, as described by Short et al., *Nuc. Acids Res.* **16**: 7583 (1988). Several clones with inserts of approximately 1.6 kilobase pairs in length were obtained from this screening procedure.

15 DNA Sequencing

20 [0061] Rescued plasmids were sequenced with Sequenase (USB, Cleveland, OH) by the dideoxy chain termination method, according to the manufacturer's protocol described by Tabor et al., *J. Biol Chem.* **214**: 6447 (1989). Nucleotide sequence analysis defined an open reading frame of 1155 nucleotides, encoding a putative protein (Dlk) of 385 amino acids with a molecular weight of 41,320 daltons. This open reading frame was classified as coding by both Fickett's and Shepherd's methods. Fickett et al., *Nucleic Acids Res.* **10**: 5303 (1982); Shepherd et al., *Meth. Enzymol.* **188**: 180 (1990). The open reading frames were identified by software programs implementing these methods (PC/Gene software package, Intelligenetics Inc. (Mountain View, CA); A. Bairoch, Ph.D thesis, University of Geneva, (1990)).

25 In vitro Translation of Dlk Polypeptide

30 [0062] *In vitro* translation assays from mouse *dlk* mRNA were performed using a rabbit reticulocyte lysate system from NEN (Boston, MA), according to manufacturer's protocol, as described by Lockhard et al., *Biochem. Biophys. Res. Comm.* **37**: 204 (1969).

35 [0063] The *d lk* mRNA was selected by hybridization of poly A+ RNA from Swiss 3T3 fibroblasts with denatured full-length *d lk* immobilized on nitrocellulose filters. (*d lk* mRNA was selected by hybridization of 2 μg of poly A+ Swiss 3T3 RNA with 5 μg of nitrocellulose-immobilized denatured *d lk*.) The bound RNA was eluted by boiling. Mouse *d lk* mRNA was also prepared *in vitro* using two different full length *d lk* cDNAs cloned in pGEM4Z (Promega). These three mRNAs were used as templates for *in vitro* translation.

40 [0064] Labelled proteins were analyzed in a 12% polyacrylamide gel followed by fluorography. A protein band of approximately 42 kilodaltons was present in all three samples, in agreement with the predicted molecular weight of *Dlk* polypeptide.

40 A Comparison between mice and humans

45 [0065] Mouse and human *d lk* polynucleotide sequences are 86.2% identical and 90.1% similarity in their amino acid sequence. They share many potential sites of biological activity, including 6 EGF-like repeats (highly homologous to those found in invertebrate neurogenic proteins) an integral transmembrane domain and a signal peptide domain.

50 [0066] The structural characteristics of *d lk* were analyzed with the program PC/Gene (Intelligenetics Inc. (Mountain View, CA), A. Bairoch, Ph.D thesis, University of Geneva (1990)). The transmembrane domain was identified using the program RAOARGOS, implementing the method of Rao and Argos, *Biochim. Biophys. Acta* **869**: 197 (1986). The signal peptide was analyzed with the program PSIGNAL, according to the method of Von Heijne, *Nucleic Acids Res.* **14**: 4683 (1986).

55 Example 2. A COMPARISON BETWEEN pPG2 and *d lk* GENE EXPRESSION IN MICE, & *d lk* GENE EXPRESSION IN HUMANS

55 [0067] In normal tissues of human, mouse and hamster origin, *d lk* expression was detected according to the present invention, only in adrenal and placental tissue. Similarly, pG2 expression was known to be restricted to adrenal glands in normal human tissues.

[0068] The *d lk* mRNA was detected by Northern analysis in human and rat pheochromocytoma (PC12) cell lines. pG2 was identified in pheochromocytoma cell lines by Helman et al., *PNAS USA* **84**: 2336 (1987).

[0069] According to the present invention, *dik* was detected in neuroblastoma (SK-N-SH) cells. pG2 expression in neuroblastoma cell lines was detected in differentiated cells, but absent from undifferentiated neuroblastoma cell lines. Cooper et al., *Cell Growth and Diff.* 1: 149 (1989).

[0070] In addition, other cells which express *dik* identified by the present invention include certain SCLC cell lines. 5 Also, murine Swiss 3T3 fibroblasts were found to express *dik* using a human *dik* probe under high stringency conditions. Balb/c 3T3 fibroblasts were found to express *dik* using a human *dik* probe under high stringency conditions. Balb/c 3T3 fibroblasts RNA were negative for *dik* expression under these conditions.

[0071] To explore the relationship between mouse *dik* and human pG2, cDNA clones were isolated from a λ gt10 human adrenal gland library (Clontech, Palo Alto, CA) using mouse *dik* as a hybridization probe. Even under low stringency conditions, no clones were isolated which encoded proteins having structural characteristics similar to those reported for pG2. The cDNA inserts from positive λ clones were subcloned into pGEM4Z (Promega, Madison WI) and sequenced according to the method of Example 1. Polynucleotide sequence data from several full-length clones indicated that these cDNAs showed a 82.1% sequence identity with mouse *dik* and encoded the human homolog of the mouse *dik* protein (Fig. 1).

[0072] Dlk's structural characterization is very different from that predicted for pG2 protein (Helman et al., *supra*. 15 (1987). The former protein consists of a 286 amino acid sequence (about 30 kDa), contains no EGF-like repeats and no signal peptide or transmembrane domains. This was so, despite a finding of an 81.2% nucleotide sequence identity of *dik* with pG2 is identified correctly as the *dik* polynucleotide molecule shown in Figure 1.

20 Example 3. *dik*/Dlk HOMOLOGY WITH OTHER GENES & PROTEINS

[0073] The *dik* nucleic acid sequence shows a high degree of homology with the EGF-like neurogenic genes of *Drosophila*, which are involved in the decisions taken by the cells of the embryonal ectoderm to differentiate into epidermal or neuronal cells. Genes which were found to have highest homology to Dlk include: Delta, Notch and Serrate 25 of *D. melanogaster*, lin-12 and glp-1 of *C. elegans*, and uEGF1 of the sea urchin. Although the degree of homology varied between the individual proteins and Dlk, regions of maximum homology exhibited up to 33% amino acid identity, which rose to around 75%, with allowance for conservative amino acid substitutions.

[0074] Figure 4 shows the alignment of mouse or human *dik* EGF-like repetitive sequences with consensus 30 sequences of EGF sequence repeats of several proteins. The alignment of the EGF-like repeats was done using the program CLUSTAL, described by Higgins et al., *Gene* 73: 237 (1988). The sites of potential biological importance were analyzed with the program PROSITE. Residues highly conserved among homeotic genes also are conserved in *dik*. This finding confirms that *dik* is a member of the family of EGF-like homeotic genes. The amino acid sequence and 35 structure of the EGF-like repeats, as well as the overall structure of *dik*, are more related to the invertebrate homeotic genes than to other vertebrate non-homeotic EGF-like proteins, such as EGF-precursor, TGF α , the α , β 1 and β 2 chains of laminin, coagulation factors, or complement proteins, previously thought to be the mammalian counterparts of the invertebrate homeotic genes.

[0075] The *dik* gene was detected in species ranging from birds to human, including: yeast, *Drosophila*, *Xenopus*, mouse, rat, rabbit, chicken, dog, cow, monkey and human. However, despite the structural homology with invertebrate proteins, the *dik* gene is absent from invertebrates and lower vertebrates.

[0076] The program PCOMPARE, described by Needleman et al., *Mol. Biol.* 48: 443 (1970), included in PC/Gene 40 was used for the analysis of homology. In this method, the optimal alignment score between two proteins were compared with the statistical distribution of 100 random alignments. An alignment score of greater than 5 positive standard deviations from the mean random alignment distribution was considered significant, particularly when no functional or structural relationship between the proteins compared is known. Representative alignment scores were determined: 45 *Delta*, 20.2; *Serrate*, 19.7; *TAN-1*, 16.2; *Notch*, 14.6; *Xotch*, 13.6; *Drosophila Laminin* β 2, 6.3; mouse *Laminin* β 2, 4.1; human coagulation factor XII, 2.8; and human EGF precursor, 0.6.

Example 4. *dik* EXPRESSION: NORTHERN BLOT mRNA ANALYSIS

[0077] Expression of *dik* was detected by Northern analysis in SCLC lines NCI-H510, NCI-H69 and NCI-N592; in 50 human neuroblastoma line SK-N-SH, and in the rat pheochromocytoma PC-12 cell line. Twenty μ g of total RNA or 2 μ g of poly A+ were separated in a 1% agarose gel and then blotted to a nitrocellulose filter (described in Example 1).

[0078] A 1.6 Kb band corresponding to *dik* was observed only in the RNA samples from the SCLC cell lines NCI- 55 N592, NCI-H69 and NCI-H510, and in Swiss 3T3 fibroblasts. Mouse Swiss 3T3 fibroblast RNA also showed a high degree of expression of *dik*, even when the hybridization was performed under high stringency conditions using human *dik* as a probe. Similar results were obtained using mouse *dik* as a probe. Balb/c 3T3 fibroblast RNA was negative for *dik* expression under these conditions. Ewing's sarcoma cell lines SK-ES-1, A4573 and TC106 did not express *dik*.

[0079] In normal tissues of mouse, hamster, and human origin *dik* expression was detected exclusively in the adre-

nal gland.

SEQUENCE LISTING

5 | GENERAL INFORMATION:

(i) APPLICANT: The Government of the United States, as Represented by the
Secretary, Department of Health and Human Services

10 (ii) TITLE OF INVENTION: DELTA-LIKE GENE EXPRESSED IN
NEUROENDOCRINE TUMORS

(iii) NUMBER OF SEQUENCES: 12

15 (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Knobbe, Martens, Olson and Bear
(B) STREET: 620 newport Center Drive, Sixteenth Floor
(C) CITY: Newport Beach
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92660

20 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

25 (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US93/12015
(B) FILING DATE: 10-DEC-1993
(C) CLASSIFICATION:

30 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (619) 235-8550
(B) TELEFAX: (619) 235-0176

| INFORMATION FOR SEQ ID NO:1:

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 385 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 Met Ile Ala Thr Gly Ala Leu Leu Arg Val Leu Leu Leu Leu Ala
 1 5 10 15
 Phe Gly His Ser Thr Tyr Gly Ala Glu Cys Asp Pro Pro Cys Asp Pro
 20 25 30
 10 Gln Tyr Gly Phe Cys Glu Ala Asp Asn Val Cys Arg Cys His Val Gly
 35 40 45
 Trp Glu Gly Pro Leu Cys Asp Lys Cys Val Thr Ala Pro Gly Cys Val
 50 55 60
 15 Asn Gly Val Cys Lys Glu Pro Trp Gln Cys Ile Cys Lys Asp Gly Trp
 65 70 75 80
 Asp Gly Lys Phe Cys Glu Ile Asp Val Arg Ala Cys Thr Ser Thr Pro
 85 90 95
 20 Cys Ala Asn Asn Gly Thr Cys Val Asp Leu Glu Lys Gly Gln Tyr Glu
 100 105 110
 25 Cys Ser Cys Thr Pro Gly Phe Ser Gly Lys Asp Cys Gln His Lys Ala
 115 120 125
 Gly Pro Cys Val Ile Asn Gly Ser Pro Cys Gln His Gly Gly Ala Cys
 130 135 140
 30 Val Asp Asp Glu Gly Gln Ala Ser His Ala Ser Cys Leu Cys Pro Pro
 145 150 155 160
 Gly Phe Ser Gly Asn Phe Cys Glu Ile Val Ala Ala Thr Asn Ser Cys
 165 170 175
 35 Thr Pro Asn Pro Cys Glu Asn Asp Gly Val Cys Thr Asp Ile Gly Gly
 180 185 190
 Asp Phe Arg Cys Arg Cys Pro Ala Gly Phe Val Asp Lys Thr Cys Ser
 195 200 205
 40 Arg Pro Val Ser Asn Cys Ala Ser Gly Pro Cys Gln Asn Gly Gly Thr
 210 215 220
 Cys Leu Gln His Thr Gln Val Ser Phe Glu Cys Leu Cys Lys Pro Pro
 225 230 235 240
 45 Phe Met Gly Pro Thr Cys Ala Lys Lys Arg Gly Ala Ser Pro Val Gln
 245 250 255
 Val Thr His Leu Pro Ser Gly Tyr Gly Leu Thr Tyr Arg Leu Thr Pro
 260 265 270
 50

Gly Val His Glu Leu Pro Val Gln Gln Pro Glu Gln His Ile Leu Lys
 275 280 285
 5 Val Ser Met Lys Glu Leu Asn Lys Ser Thr Pro Leu Leu Thr Glu Gly
 290 295 300
 Gln Ala Ile Cys Phe Thr Ile Leu Gly Val Leu Thr Ser Leu Val Val
 305 310 315 320
 10 Leu Gly Thr Val Ala Ile Val Phe Leu Asn Lys Cys Glu Thr Trp Val
 325 330 335
 Ser Asn Leu Arg Tyr Asn His Thr Phe Arg Lys Lys Asn Leu Leu
 340 345 350
 15 Leu Gln Tyr Asn Ser Gly Glu Glu Leu Ala Val Asn Ile Ile Phe Pro
 355 360 365
 Glu Lys Ile Asp Met Thr Thr Phe Asn Lys Glu Ala Gly Asp Glu Glu
 20 370 375 380
 25 Ile
 385

2) INFORMATION FOR SEQ ID NO:2:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 383 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 30
 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: NO
 35 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

40 Met Thr Ala Thr Glu Ala Leu Leu Arg Val Leu Leu Leu Leu Ala
 1 5 10 15
 Phe Gly His Ser Thr Tyr Gly Ala Glu Cys Phe Pro Ala Cys Asn Pro
 20 25 30
 45 Gln Asn Gly Phe Cys Glu Asp Asp Asn Val Cys Arg Cys Gln Pro Gly
 35 40 45
 Trp Gln Gly Pro Leu Cys Asp Gln Cys Val Thr Ser Pro Gly Cys Leu

50

55

	50	55	60														
5	His	Gly	Leu	Cys	Gly	Glu	Pro	Gly	Gln	Cys	Ile	Cys	Thr	Asp	Gly	Trp	
	65				70					75						80	
	Asp	Gly	Glu	Leu	Cys	Asp	Arg	Asp	Val	Arg	Ala	Cys	Ser	Ser	Ala	Pro	
					85				90						95		
10	Cys	Ala	Asn	Asn	Gly	Thr	Cys	Val	Ser	Leu	Asp	Asp	Gly	Leu	Tyr	Glu	
					100				105					110			
	Cys	Ser	Cys	Ala	Pro	Gly	Tyr	Ser	Gly	Lys	Asp	Cys	Gln	Lys	Lys	Asp	
					115			120					125				
15	Gly	Pro	Cys	Val	Ile	Asn	Gly	Ser	Pro	Cys	Gln	His	Gly	Gly	Thr	Cys	
					130			135				140					
	Val	Asp	Asp	Glu	Gly	Arg	Ala	Ser	His	Ala	Ser	Cys	Leu	Cys	Pro	Pro	
					145			150			155			160			
20	Gly	Phe	Ser	Gly	Asn	Phe	Cys	Glu	Ile	Val	Ala	Asn	Ser	Cys	Thr	Pro	
					165				170					175			
	Asn	Pro	Cys	Glu	Asn	Asp	Gly	Val	Cys	Thr	Asp	Ile	Gly	Gly	Asp	Phe	
25					180				185				190				
	Arg	Cys	Arg	Cys	Pro	Ala	Gly	Phe	Ile	Asp	Lys	Thr	Cys	Ser	Arg	Pro	
					195				200			205					
30	Val	Thr	Asn	Cys	Ala	Ser	Ser	Pro	Cys	Gln	Asn	Gly	Gly	Thr	Cys	Leu	
					210				215			220					
	Gln	His	Thr	Gln	Val	Ser	Tyr	Glu	Cys	Leu	Cys	Lys	Pro	Glu	Phe	Thr	
					225			230			235			240			
35	Gly	Leu	Thr	Cys	Val	Lys	Lys	Arg	Ala	Leu	Ser	Pro	Gln	Gln	Val	Thr	
					245				250			255					
	Arg	Leu	Pro	Ser	Gly	Tyr	Gly	Leu	Ala	Tyr	Arg	Leu	Thr	Pro	Gly	Val	
					260				265			270					
40	His	Glu	Leu	Pro	Val	Gln	Gln	Pro	Glu	His	Arg	Ile	Leu	Lys	Val	Ser	
					275				280			285					
	Met	Lys	Glu	Leu	Asn	Lys	Lys	Thr	Pro	Leu	Leu	Thr	Glu	Gly	Gln	Ala	
					290			295				300					
45	Ile	Cys	Phe	Thr	Ile	Leu	Gly	Val	Leu	Thr	Ser	Leu	Val	Val	Leu	Gly	
					305			310			315			320			
	Thr	Val	Gly	Ile	Val	Phe	Leu	Asn	Lys	Cys	Glu	Thr	Trp	Val	Ser	Asn	
50						325				330				335			

Leu Arg Tyr Asn His Met Leu Arg Lys Lys Lys Asn Leu Leu Gln
 340 345 350

5 Tyr Asn Ser Gly Glu Asp Leu Ala Val Asn Ile Ile Phe Pro Glu Lys
 355 360 365

Ile Asp Met Thr Thr Phe Ser Lys Glu Ala Gly Asp Glu Glu Ile
 370 375 380

10 2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1556 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

25	CTAAAGGAG GTGGAGAGCG CACCGCAGCC CGGTGCAGCC CGGTGCAGCC CTGGCTTTCC	60
	CTCGCTGCG GCCCCTGCCC CCTTTCGCGT CCGCAACCAAG AAGCCCAGTG CGGCAGCCAGG	120
	GCGGGACCC GCGCCCGCAC CGCTCCCGGG ACCGCGACCC CGGCGGCCCA GAGATGACCG	180
30	GACCGAAGC CCTCCTGCGC GTCCTCTTGC TCCTGCTGGC TTTCGGCCAC AGCACCTATG	240
	GGCTGAATG CTTCCCGGCC TGCAACCCCC AAAATGGATT CTGCGAGGAT GACAATGTTT	300
	CAGGTGCCA GCCTGGCTGG CAGGGTCCCC TTTGTGACCA GTGCGTGACC TCTCCGGCT	360
35	CCTTCACGG ACTCTGTGGA GAACCCGGGC AGTGCATTG CACCGACGGC TGGGACGGGG	420
	GCTCTGTGA TAGAGATGTT CGGGCCTGCT CCTCGCCCC CTGTGCCAAC AACGGGACCT	480
	CGTGAGCCT GGACGATGGC CTCTATGAAT GCTCCTGTGC CCCCGGGTAC TCAGGGAAAGG	540
40	CTGCCAGAA AAAGGACGGG CCCTGTGTGA TCAACGGCTC CCCCTGCCAG CACGGAGGCA	600
	CTGCGTGGA TGATGAGGGC CGGGCCTCCC ATGCCTCCTG CCTGTGCCCC CCTGGCTTCT	660
	AGGCAATTG CTGCGAGATC GTGGCCAACA GCTGCACCCCC CAACCCATGC GAGAACGACG	720
45	CGTCTGCAC TGACATTGGG GGCGACTTCC GCTGCCGGTG CCCAGCCGGC TTCAATCGACA	780

5	ACCTGCAG CCGCCCGGTG ACCAACTGCG CCAGCAGCCC GTGCCAGAAC GGGGGCACCT	840
	CTGCAGCA CACCCAGGTG AGCTACGAGT GTCTGTGCAA GCCCCAGTGC ACAGGTCTCA	900
	CTGTCAA GAAGCGCGCG CTGAGCCCC AGCAGGTAC CCGTCTGCC AGCGGCTATG	960
	CTGGCCTA CCGCCTGACC CCTGGGGTGC ACGAGCTGCC GGTGCAGCAG CCGGAGCACC	1020
10	ATCCTGAA GGTGTCCATG AAAGAGCTCA ACAAGAAAAC CCCTCTCCTC ACCGAGGGCC	1080
	GCCATCTG CTTCACCATC CTGGCGTGC TCACCAGCCT GGTGGTGCTG GGCACTGTGG	1140
	ATCGTCTT CCTCAACAAG TGCGAGACCT GGGTGTCCAA CCTGCCTAC AACCACATGC	1200
15	CGGAAGAA GAAGAACCTG CTGCTTCAGT ACAACAGCGG GGAGGACCTG GCCGTCAACA	1260
	ATCTTCCC CGAGAAGATC GACATGACCA CCTTCAGCAA GGAGGCCGGC GACGAGGAGA	1320
	TAAGCAGC GTTCCCACAG CCCCCCTCTAG ATTCTTGGAG TTCCGCAGAG CTTACTATAC	1380
20	GGTCTGTC CTAATCTTGT TGGTGTTCGC TATCTTTGT GTCAAATCTG GTGAACGCTA	1440
	TTACATA TATTGTCTTT GTGCTGCTGT GTGACAAACG CAATGCAAAA ACAATCCTCT	1500
	TCTCTCT TAATGCATGA TACAGAATAA TAATAAGAAT TTCATCTTTA AATGAG	1556

25 INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1573 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

30 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

40	TGCAACCC TAGCTTTCTT CCCGCTGGAC GCCCCGTGCC CTTCTGTGGT CCGCAACCAG	60
	GCCCCAGCG CACGCCCGG AGCAGCCCT GCACCGCCTC CGCTCCCCGG ACCGCGACCC	120
	GCCGCCCG GAGATGATCG CGACCGGAGC CCTCCTGCCTC GTCTCTTGC TCCTGCTGGC	180
45	TCGGCCAC AGCACCTATG GGGCTGAATG CGACCCACCC TGTGACCCCC AGTATGGATT	240

5	'GCGAGGCT GACAATGTCT GCAGGTGCCA TGTGGCTGG GAGGGTCCCC TCTGTGACAA	300
	'GTGTAACT GCCCCTGGCT GTGTCAATGG AGTCTGCAAG GAACCATGGC AGTGCATCTG	360
	AAGGATGGC TGGGACGGGA AATTCTGCGA AATAGACGTT CGGGCTTGCA CCTCAACCCC	420
	'GCGCCAAC AATGGAACCTT GCGTGGACCT GGAGAAAGGC CAGTACGAAT GCTCCTGCAC	480
10	CCTGGGTTTC TCTGGAAAGG ACTGCCAGCA CAAGGCTGGG CCCTGCGTGA TCAATGGTT	540
	CCCTGCCAG CACGGAGGCG CCTGCGTGGA TGATGAGGGC CAGGCCTCGC ATGCTTCCTG	600
	CTGTGCCCC CCTGGCTTCT CAGGCAACTT CTGTGAGATC GTAGCCGCAA CCAACAGCTG	660
15	ACCCCTAAC CCATGCGAGA ACGATGGCGT CTGCACCGAC ATCGGGGGTG ACTTCCGTTG	720
	CGCTGCCCA GCTGGATTG TCGACAAGAC CTGCAGCCGC CCGGTGAGCA ACTGCCAG	780
	GGCCCGTGC CAGAACGGGG GCACCTGCCT CCAGCACACC CAGGTGAGCT TCGAGTGTCT	840
20	'GCAAGCCC CCGTTCATGG GTCCCACGTG CGCGAAGAAG CGCGGGGCTA GCCCGTGCA	900
	GTCACCCAC CTGCCAGCG GCTATGGGCT CACCTACCGC CTGACCCCCCG GGGTGCACGA	960
	CTGCCTGTT CAGCAGCCCCG AGCAACACAT CCTGAAGGTG TCCATGAAAG AGCTAACAA	1020
25	AGTACCCCT CTCCTCACCG AGGGACAGGC CATCTGCTTC ACCATCCTGG GCGTGCAC	1080
	AGCCTGGTG GTGCTGGCA CCGTGGCCAT CGTCTTCTC AACAAAGTGG AAACCTGGGT	1140
	TCCAACCTG CGCTACAACC ACACGTTCG CAAGAAGAAG AACCTCCTGT TGCAGTATAA	1200
30	AGCGGGAG GAGCTGGCGG TCAATATCAT CTTCCCGAG AAGATTGACA TGACCACCTT	1260
	AACAAGGAG GCTGGTGATG AGGAGATCTA AGCAGCGTTC CCCACCCCCA CTCCCAGGCC	1320
	TCACCCCCG ACCCCGACCC AGGCCCTCTC TATTACCGGG TTCCTTCTAGA GCTCTCTACC	1380
35	AGTCTGGCT TTTTGTGGTG GAGTTTGCTC TATTGTGTGG AATCGAGTGA AGCCTATGCT	1440
	ACATATATT GTCTTGTGTT GCTGTGTGCC ATGCTACCTC GCTATCTAAG AACCCCTTCC	1500
	CCCTATTAA TGCATGATAA TGAATAATAA TAATAAGAAT TTCATCTCTA AATGAAAAAA	1560
40	AAAAAAAAA AAG	1573

2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single

5 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (v) FRAGMENT TYPE: internal

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

30 Cys Pro Cys Asn Gly Gly Cys Asp Tyr Cys Cys Pro Gly Phe Gly Cys
1 5 10 15

35 2) INFORMATION FOR SEQ ID NO:6:

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: peptide

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

60 Cys Pro Cys Asn Gly Gly Thr Cys Tyr Cys Cys Gly Phe Gly Cys
1 5 10 15

65 2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Cys Pro Cys Asn Gly Gly Thr Cys Asp Phe Cys Cys Gly Gly Cys
1 5 10 15

INFORMATION FOR SEQ ID NO:8:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys Pro Cys Asn Gly Gly Thr Cys Asp Tyr Cys Cys Gly Phe Gly Cys
1 5 10 15

35 INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

45 (iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

5 Cys Ser Pro Cys Asn Gly Gly Thr Cys Asp Tyr Cys Cys Gly Phe Gly
1 5 10 15

Cys

10 INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Cys Pro Cys Asn Gly Gly Thr Cys Asp Tyr Cys Cys Gly Phe Gly Cys
1 5 10 15

30) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
35 (B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

50

55

Cys Leu Cys Asn Gly Cys Ile Tyr Cys Cys Gly Tyr Gly Cys
 1 5 10

5) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (v) FRAGMENT TYPE: internal

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys Pro Ser Ser Tyr Asp Gly Tyr Cys Leu Asn Gly Gly Val Cys Met
 1 5 10 15

His Ile Glu Ser Asp Leu Ser Tyr Thr Cys Asn Cys Val Ile Gly Tyr
 20 25 30

30 Ser Gly Asp Arg Cys
 35

35

Claims

40 1. An isolated Dlk polypeptide consisting essentially of the amino acid sequence shown in Figure 1B.

2. An isolated polynucleotide molecule which encodes a human Dlk polypeptide according to Claim 1.

3. An isolated polynucleotide molecule consisting essentially of the polynucleotide sequence shown in Figure 2.

45 4. An isolated polynucleotide molecule which encodes a murine Dlk polypeptide, wherein said Dlk polypeptide comprises the amino acid sequence shown in Figure 1A.

5. A method for detecting a tumor which expresses *dlk*, comprising the steps of

50 (a) contacting a sample suspected of being tumorigenic with *dlk* polynucleotide molecule, under conditions permissive of hybridization between *dlk* polynucleotide molecule and said sample and
 (b) detecting the presence of hybridization between said polynucleotide molecule and said sample.

55 6. A method according to Claim 5, wherein said *dlk* polynucleotide molecule is an isolated polynucleotide molecule according to Claim 3.

7. A method according to Claim 5 for detecting a tumor identified as small cell lung carcinoma, wherein, in step (a),

said sample comprises bronchial epithelial cells.

5

10

15

20

25

30

35

40

45

50

55

Figure 1A
Figure 1B

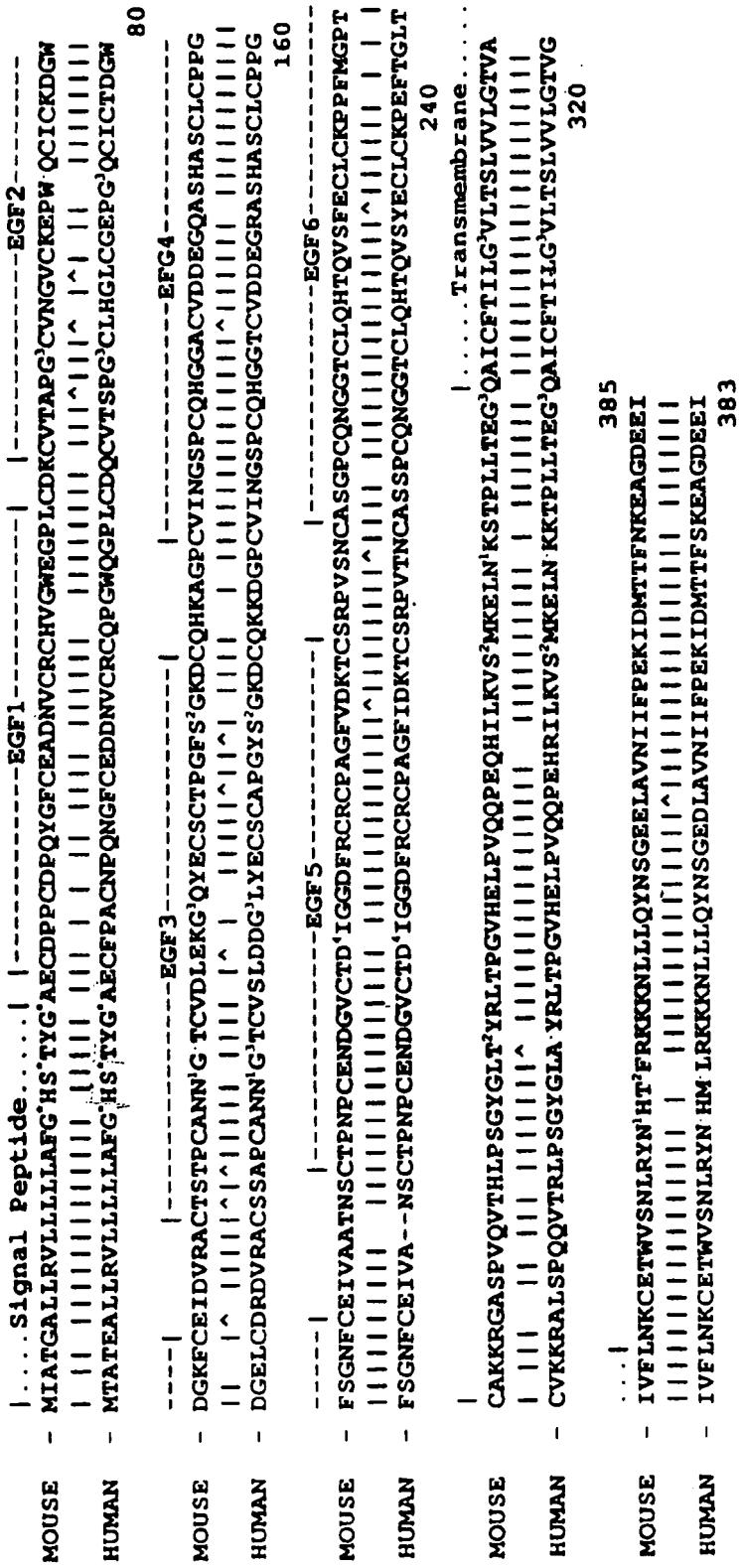


Figure 2

	3	9	15	21	27	33	39	45							
1	TCT	AAA	GGA	GGT	GGA	GAG	CGC	ACC	GCA	GCC	CGG	TGC	AGC	CCG	GTG
46	CAG	CCC	TGG	CTT	TCC	CCT	CGC	TGC	GGC	CCG	TGC	CCC	CTT	TCC	CGT
91	CCG	CAA	CCA	GAA	GCC	CAG	TGC	GGC	GCC	AGG	AGC	CGG	ACC	CGC	GCC
136	CGC	ACC	GCT	CCC	GGG	ACC	GCG	ACC	CCG	GCC	GCC	CAG	AGA	TGA	CCG
181	CGA	CCG	AAG	CCC	TCC	TGC	GCG	TCC	TCT	TGC	TCC	TGC	TGG	CTT	TGG
226	GCC	ACA	GCA	CCT	ATG	GGG	CTG	AAT	GCT	TCC	CGG	CCT	GCA	ACC	CCC
271	AAA	ATG	GAT	TCT	GGC	AGG	ATG	ACA	ATG	TTT	GCA	GGT	GCC	AGC	CTG
316	GCT	GGC	AGG	GTC	CCC	TTT	GTG	ACC	AGT	GCG	TGA	CCT	CTC	CCG	GCT
361	GCC	TTC	ACG	GAC	TCT	GTG	GAG	AAC	CCG	GCG	AGT	GCA	TTT	GCA	CCG
406	ACG	GCT	GGG	ACG	GGG	AGC	TCT	GTG	ATA	GAG	ATG	TTC	GGG	CCT	GCT
451	CCT	CGG	CCC	CCT	GTG	CCA	ACA	ACG	GGA	CCT	GCG	TGA	GCC	TGG	ACG
496	ATG	GCC	TCT	ATG	AAT	GCT	CCT	GTG	CCC	CCG	GGT	ACT	CGG	GAA	AGG
541	ACT	GCC	AGA	AAA	AGG	ACG	GGC	CCT	GTG	TGA	TCA	ACG	GCT	CCC	CCT
586	GCC	AGC	ACG	GAG	GCA	CCT	GCG	TGG	ATG	ATG	AGG	GCC	GGG	CCT	CCC
631	ATG	CCT	CCT	GCC	TGT	GCC	CCC	CTG	GCT	TCT	CAG	GCA	ATT	TCT	GGG
676	AGA	TCG	TGG	CCA	ACA	GCT	GCA	CCC	CCA	ACC	CAT	GCG	AGA	ACG	ACG
721	GCG	TCT	GCA	CTG	ACA	TTG	GGG	GCG	ACT	TCC	GCT	GCC	GGT	GCC	CAG
766	CCG	GCT	TCA	TCG	ACA	AGA	CCT	GCA	GCC	GCC	CGG	TGA	CCA	ACT	GCG
811	CCA	GCA	GCC	CGT	GCC	AGA	ACG	GGG	GCA	CCT	GCC	TGC	AGC	ACA	CCC
856	AGG	TGA	GCT	ACG	AGT	GTC	TGT	GCA	AGC	CCG	AGT	TCA	CAG	GTC	TCA
901	CCT	GTG	TCA	AGA	AGC	GCG	CAC	TGA	GCC	CCC	AGC	AGG	TCA	CCC	GTC
946	TGC	CCA	GCG	GCT	ATG	GGC	TGG	CCT	ACC	GCC	TGA	CCC	CTG	GGG	TGC
991	ACG	AGC	TGC	CGG	TGC	AGC	AGC	CGG	AGC	ACC	GCA	TCC	TGA	AGG	TGT
1036	CCA	TGA	AAG	AGC	TCA	ACA	AGA	AAA	CCC	CTC	TCC	TCA	CCG	AGG	GCC
1081	AGG	CCA	TCT	GCT	TCA	CCA	TCC	TGG	GCG	TGC	TCA	CCA	GCC	TGG	TGG
1126	TGC	TGG	GCA	CTG	TGG	GTA	TGC	TCT	TCC	TCA	ACA	AGT	GCG	AGA	CCT
1171	GGG	TGT	CCA	ACC	TGC	GCT	ACA	ACC	ACA	TGC	TGC	GGA	AGA	AGA	AGA
1216	ACC	TGC	TGC	TTC	AGT	ACA	ACA	GCG	GGG	AGG	ACC	TGG	CCG	TCA	ACA
1261	TCA	TCT	TCC	CCG	AGA	AGA	TGC	ACA	TGA	CCA	CCT	TCA	GCA	AGG	AGG
1306	CCG	GCG	ACG	AGG	AGA	TCT	AAG	CAG	CGT	TCC	CAC	AGC	CCC	CTC	TAG
1351	ATT	CTT	GGA	GTT	CCG	CAG	AGC	TTA	CTA	TAC	GCG	GTC	TGT	CCT	AAT
1396	CTT	TGT	GGT	GTT	CGC	TAT	CTC	TTG	TGT	CAA	ATC	TGG	TGA	ACG	CTA
1441	CGC	TTA	CAT	ATA	TTG	TCT	TTG	TGC	TGC	TGT	GTG	ACA	AAC	GCA	ATG
1486	CAA	AAA	CAA	TCC	TCT	TTG	TCT	CTC	TTA	ATG	CAT	GAT	ACA	GAA	TAA
1531	TAA	TAA	GAA	TTT	CAT	CTT	TAA	ATG	AG						

Total number of bases is: 1556.

Figure 3

	3	9	15	21	27	33	39	45							
1	GGT	GCA	ACC	CTA	GCT	TTC	TTC	CCG	CTG	GAC	GCC	CGT	GCC	CCC	TTC
46	GTC	CGC	AAC	CAG	AAG	CCC	AGC	GCA	CGC	CCC	GGA	GCA	GGC	CCT	
91	GCA	CCG	CCT	CCG	CTC	CCC	GGA	CCG	CGA	CCC	AGG	CCG	CCC	CGA	AT
136	GAT	CGC	GAC	CGG	AGC	CCT	CCT	GCG	CGT	CCT	CTT	GCT	CCT	GCT	GGC
181	TTT	CGG	CCA	CAG	CAC	CTA	TGG	GGC	TGA	ATG	CGA	CCC	ACC	CTG	TGA
226	CCC	CCA	GTA	TGG	ATT	CTG	CGA	GGC	TGA	CAA	TGT	CTG	CAG	GTG	CCA
271	TGT	TGG	CTG	GGA	GGG	TCC	CCT	CTG	TGA	CAA	GTG	TGT	AAC	TGC	CCC
316	TGG	CTG	TGT	CAA	TGG	AGT	CTG	CAA	GGA	ACC	ATG	GCA	GTG	CAT	CTG
361	CAA	GGA	TGG	CTG	GGA	CGG	GAA	ATT	CTG	CGA	AAT	AGA	CGT	TCG	GGC
406	TTG	CAC	CTC	AAC	CCC	CTG	CGC	CAA	CAA	TGG	AAC	TTG	CGT	GGA	CCT
451	GGA	GAA	AGG	CCA	GTA	CGA	ATG	CTC	CTG	CAC	ACC	TGG	GTT	CTC	TGG
496	AAA	GGA	CTG	CCA	GCA	CAA	GGC	TGG	GCC	CTG	CGT	GAT	CAA	TGG	TTC
541	TCC	CTG	CCA	GCA	CGG	AGG	CGC	CTG	CGT	GGA	TGA	TGA	GGG	CCA	GGC
586	CTC	GCA	TGC	TTC	CTG	CCT	GTG	CCC	CCC	TGG	CTT	CTC	AGG	CAA	CTT
631	CTG	TGA	GAT	CGT	AGC	CGC	XAC	CAA	CAG	CTG	TAC	CCC	TAA	CCC	ATG
676	CGA	GAA	CGA	TGG	CGT	CTG	CAC	CGA	CAT	CGG	GGG	TGA	CTT	CCG	TTG
721	CCG	CTG	CCC	AGC	TGG	ATT	CGT	CGA	CAA	GAC	CTG	CAG	CCG	CCC	GGT
766	GAG	CAA	CTG	CGC	CAG	TGG	CCC	GTG	CCA	GAA	CGG	GGG	CAC	CTG	CCT
811	CCA	GCA	CAC	CCA	GGT	GAG	CTT	CGA	GTG	TCT	GTG	CAA	GCC	CCC	GTT
856	CAT	GGG	TCC	CAC	GTG	CGC	GAA	GAA	GCG	CGG	GGC	TAG	CCC	CGT	GCA
901	GGT	CAC	CCA	CCT	GCC	CAG	CGG	CTA	TGG	GCT	CAC	CTA	CCG	CCT	GAC
946	CCC	CGG	GGT	GCA	CGA	GCT	GCC	TGT	TCA	GCA	GCC	CGA	GCA	ACA	CAT
991	CCT	GAA	GGT	GTC	CAT	GAA	AGA	GCT	CAA	CAA	GAG	TAC	CCC	TCT	CCT
1036	CAC	CGA	GGG	ACA	GGC	CAT	CTG	CTT	CAC	CAT	CCT	GGG	CGT	GCT	CAC
1081	CAG	CCT	GGT	GGT	GCT	GGG	CAC	CGT	GGC	CAT	CGT	CTT	TCT	CAA	CAA
1126	GTG	CGA	AAC	CTG	GGT	GTC	CAA	CCT	CGG	CTA	CAA	CCA	CAC	GTT	TCG
1171	CAA	GAA	GAA	GAA	CCT	CCT	GTT	GCA	GTA	TAA	CAG	CGG	CGA	GGA	GCT
1216	GGC	GGT	CAA	TAT	CAT	CTT	CCC	CGA	GAA	GAT	TGA	CAT	GAC	CAC	CTT
1261	CAA	CAA	GGA	GGC	TGG	TGA	TGA	GGA	GAT	CTA	AGC	AGC	GTT	CCC	CAC
1306	CCC	CAC	TCC	CAG	GCC	CTT	CAC	CCC	GAC	CCC	GAC	CCA	GGC	CCT	CTC
1351	TAT	TAC	CGG	GGT	CCT	TTA	GAG	CTC	TCT	ACC	GAG	TCT	GGC	TTT	TTG
1396	TGG	TGG	AGT	TTG	CTC	TAT	TGT	GTG	GAA	TGG	AGT	GAA	GCC	TAT	GCT
1441	TAC	ATA	TAT	TGT	CTT	GTG	TTG	CTG	TGT	GCC	ATG	CTA	CCT	CGC	TAT
1486	CTA	AGA	ACC	CCT	TCC	CTA	TTA	ATG	CAT	GAT	AAT	GAA	TAA	TAA	
1531	TAA	TAA	GAA	TTT	CAT	CTC	TAA	ATG	AAA	AAA	AAA	AAA	AAA	AAA	G

Total number of bases is: 1573.

Consensus

dlk	C	PC	NGG	C	D	Y	C	C	PGF	G	C
Delta	C	PC	NGGTC			Y	C	C	GF	G	C
Serrate	C	PC	NGGTC	D		F	C	C	G	G	C
TAN-1	C	PC	NGGTC	D		Y	C	C	GF	G	C
Notch	C	S	PC	NGGTC	D	Y	C	C	GF	G	C
Kotch	C	PC	NGGTC	D		Y	C	C	GF	G	C
Lin-12	CL		C	N	G	CI	Y	C	C	GY	G

EGF mouse CPSSYDGYCLNGGVCMHIESDLSYT CNCVIGYSGDRC

FIGURE 4